

DOCKET NO.: 273943US90PCT

0/539609
JC17 Rec'd PCT/PTO 17 JUN 2005

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN RE APPLICATION OF: Shoko ARAKI, et al.

SERIAL NO.: NEW U.S. PCT APPLICATION

FILED: HERewith

INTERNATIONAL APPLICATION NO.: PCT/JP04/12629

INTERNATIONAL FILING DATE: September 1, 2004

FOR: SIGNAL SEPARATION METHOD, SIGNAL SEPARATION DEVICE, SIGNAL SEPARATION PROGRAM AND RECORDING MEDIUM

REQUEST FOR PRIORITY UNDER 35 U.S.C. 119
AND THE INTERNATIONAL CONVENTION

Commissioner for Patents
Alexandria, Virginia 22313

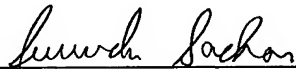
Sir:

In the matter of the above-identified application for patent, notice is hereby given that the applicant claims as priority:

<u>COUNTRY</u>	<u>APPLICATION NO</u>	<u>DAY/MONTH/YEAR</u>
Japan	2003-309720	02 September 2003
Japan	2004-195818	01 July 2004
Japan	2004-195867	01 July 2004

Certified copies of the corresponding Convention application(s) were submitted to the International Bureau in PCT Application No. PCT/JP04/12629.

Respectfully submitted,
OBLON, SPIVAK, McCLELLAND,
MAIER & NEUSTADT, P.C.



Masayasu Mori
Attorney of Record
Registration No. 47,301
Surinder Sachar
Registration No. 34,423

Customer Number

22850

(703) 413-3000
Fax No. (703) 413-2220
(OSMMN 08/03)

U / 539609

PCT/JP2004/012629

日本国特許庁
JAPAN PATENT OFFICE

03.09.2004

17 JUN 2005

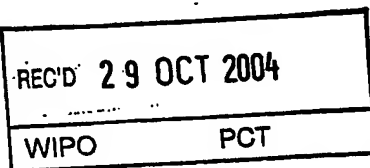
別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
Date of Application: 2003年 9月 2日

出願番号
Application Number: 特願2003-309720
[ST. 10/C]: [JP2003-309720]

出願人
Applicant(s): 日本電信電話株式会社

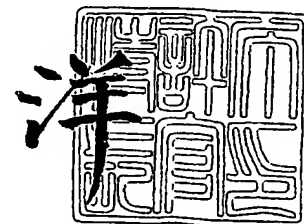


PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

2004年10月14日

特許庁長官
Commissioner,
Japan Patent Office

小川



出証番号 出証特2004-3092250

【書類名】 特許願
【整理番号】 NTTH155782
【提出日】 平成15年 9月 2日
【あて先】 特許庁長官殿
【国際特許分類】 H03M
【発明者】
 【住所又は居所】 東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
 【氏名】 荒木 章子
【発明者】
 【住所又は居所】 東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
 【氏名】 牧野 昭二
【発明者】
 【住所又は居所】 東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
 【氏名】 向井 良
【発明者】
 【住所又は居所】 東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
 【氏名】 澤田 宏
【特許出願人】
 【識別番号】 000004226
 【氏名又は名称】 日本電信電話株式会社
【代理人】
 【識別番号】 100066153
 【弁理士】
 【氏名又は名称】 草野 卓
【選任した代理人】
 【識別番号】 100100642
 【弁理士】
 【氏名又は名称】 稲垣 稔
【手数料の表示】
 【予納台帳番号】 002897
 【納付金額】 21,000円
【提出物件の目録】
 【物件名】 特許請求の範囲 1
 【物件名】 明細書 1
 【物件名】 図面 1
 【物件名】 要約書 1
 【包括委任状番号】 9806848

【書類名】 特許請求の範囲**【請求項 1】**

N個の信号源から発せられた信号が、M ($M < N$) 個のセンサで混合して観測される状況において、当該センサによる観測結果から所定の信号を分離するブラインド信号分離装置であって、

上記センサにおいて観測された観測信号を、周波数領域の信号に変換する周波数領域変換手段と、

上記周波数領域の信号を用い、センサ間における観測値の相対値を、各周波数について算出する相対値算出手段と、

上記相対値の分布からN個の分布の平均値を算出する平均値算出手段と、

上記平均値を用い、上記周波数領域の信号から、2個以上M個以下の信号源から発せられた信号によって構成される混合信号（以下「限定混合信号」という。）を抽出するためのバイナリマスクを作成するバイナリマスク作成手段と、

上記バイナリマスクを用い、上記周波数領域の信号から、上記限定混合信号を抽出する限定混合信号抽出手段と、

上記限定混合信号をそれぞれの信号に分離し分離信号を得る信号分離手段と、

上記分離信号を時間領域の信号に戻す時間領域変換手段と、

を有することを特徴とするブラインド信号分離装置。

【請求項 2】

上記限定混合信号は、

M個の信号源から発せられた信号によって構成される混合信号である、

ことを特徴とする請求項 1 記載のブラインド信号分離装置。

【請求項 3】

上記バイナリマスクは、

第 1 の上記平均値と第 2 の上記平均値を含む所定の範囲内にある上記相対値に対して 0 以外の数値を出力し、この範囲内にない上記相対値に対して 0 を出力する関数である、

ことを特徴とする請求項 1 或いは 2 に記載のブラインド信号分離装置。

【請求項 4】

上記第 1 の平均値から上記第 2 の平均値までの範囲に含まれる上記平均値の数は、上記センサの数M以下であり、

上記限定混合信号抽出手段は、

上記周波数領域の信号に、上記バイナリマスクを乗じた値を、上記限定混合信号とする

ことを特徴とする請求項 3 記載のブラインド信号分離装置。

【請求項 5】

上記第 1 の平均値から上記第 2 の平均値までの範囲に含まれる上記平均値の数は、上記センサの数Mであること、

を特徴とする請求項 4 記載のブラインド信号分離装置。

【請求項 6】

上記限定混合信号抽出手段は、

上記周波数領域の信号から、この周波数領域の信号に上記バイナリマスクを乗じた値を減じることにより、上記限定混合信号を抽出する、

ことを特徴とする請求項 1 から 3 の何れかに記載のブラインド信号分離装置。

【請求項 7】

上記バイナリマスク作成手段は、

複数通りの上記バイナリマスクを作成し、

上記限定混合信号抽出手段は、

これら複数の上記バイナリマスクを用い、上記周波数領域の信号から、複数種類の上記限定混合信号を抽出する、

ことを特徴とする請求項 1 から 6 の何れかに記載のブラインド信号分離装置。

【請求項 8】

上記バイナリマスク作成手段は、

第 1 の上記平均値と第 2 の上記平均値を含む所定の範囲内にある上記相対値に対して 0 以外の数値を出力し、この範囲内にない上記相対値に対して 0 を出力する関数を、上記バイナリマスクとして作成し、

この第 1 の平均値及び第 2 の平均値となる上記平均値の組み合わせを変更することにより、複数通りの上記バイナリマスクを作成する、

ことを特徴とする請求項 7 記載のブラインド信号分離装置。

【請求項 9】

上記バイナリマスク作成手段は、

上記第 1 の平均値と上記第 2 の平均値との間に含まれない平均値を含むよう、新たな第 1 の平均値或いは第 2 の平均値を設定して選択していくことにより、複数通りの上記バイナリマスクを作成する、

ことを特徴とする請求項 8 記載のブラインド信号分離装置。

【請求項 10】

上記相対値は、

上記観測信号のセンサ間位相差及び振幅比の少なくとも一方である、

ことを特徴とする請求項 1 から 9 の何れかに記載のブラインド信号分離装置。

【請求項 11】

N 個の信号源から発せられた信号が、M ($M < N$) 個のセンサで混合して観測される状況において、当該センサによる観測結果から所定の信号を分離するブラインド信号分離方法であって、

上記センサで観測した観測信号を、周波数領域の信号に変換し、

上記周波数領域の信号を用い、センサ間における観測値の相対値を、各周波数について算出し、

上記相対値の分布から N 個の分布の平均値を算出し、

上記平均値を用い、上記周波数領域の信号から 2 個以上 M 個以下の信号源から発せられた信号によって構成される混合信号（以下「限定混合信号」という。）を抽出するためのバイナリマスクを作成し、

上記バイナリマスクを用い、上記周波数領域の信号から上記限定混合信号を抽出し、

上記限定混合信号をそれぞれの信号に分離して分離信号を得、

上記分離信号を時間領域の信号に戻すこと、

を特徴とするブラインド信号分離方法。

【請求項 12】

請求項 1 から 10 の何れかに記載されたブラインド信号分離装置としてコンピュータを機能させるためのブラインド信号分離プログラム。

【書類名】明細書

【発明の名称】ブラインド信号分離装置、ブラインド信号分離方法及びブラインド信号分離プログラム

【技術分野】

【0001】

この発明は、信号処理の技術分野に関し、特に、必要である原信号（目的信号）のみを直接観測することができず、目的信号に他の信号が重畳されて観測されるという状況において目的信号を推定する、ブラインド信号分離装置、ブラインド信号分離方法及びこれらの機能をコンピュータ上で実現するためのブラインド信号分離プログラムに関する。

【背景技術】

【0002】

従来から、複数の原信号（音声信号等）が混合された混合信号を用い、原信号や混合過程の知識を用いることなく、混合前の原信号を分離・抽出するブラインド信号分離（BSS: Blind Source Separation）技術が知られている。

図8の（a）は、このブラインド信号分離技術を概念的に例示したブロック図である。

この図に例示するように、ブラインド信号分離では、複数（この例ではN個）の信号源201から発せられた原信号 s_1, \dots, s_N が混合し、複数（この例ではM個）のセンサ202で観測される状況下において、この観測結果のみから、原信号と推測される分離信号 y_1, \dots, y_N が取り出される。ここで、信号源201から発せられた原信号 s_1, \dots, s_N が混合し、センサ202で観測されるまでの過程を「混合過程」と呼び、センサ202の観測結果から分離信号を取り出す過程を「分離過程」と呼ぶ。以下、この混合過程と分離過程をモデル化する。

【0003】

〔混合過程のモデル化〕

まず、混合過程についてモデル化する。

Nを信号源201の個数、Mをセンサ202の個数、 s_i をi番目（ $i=1, \dots, N$ ）の信号源201（以下「信号源i」という。）から発せられた信号、 h_{ji} を信号源iからj番目（ $j=1, \dots, M$ ）のセンサ202（以下「センサj」という。）までのインパルス応答とする。この際、センサjで観測される信号 x_j は、 s_i と h_{ji} の畳み込み混合

【数1】

$$x_j(t) = \sum_{i=1}^N \sum_{p=1}^P h_{ji}(p) s_i(t-p+1) \cdots (1)$$

でモデル化される。ここで「畳み込み」とは、信号の伝搬過程で、信号が遅延され、所定の係数が乗算された後、加算されることをいう。また、Pはインパルス応答の長さを、tはサンプリング時刻を、pは加算演算のための変数を、それぞれ示している。なお、N個の信号源201は統計的に互いに独立であり、それぞれの信号は十分スパースであると仮定する。また、「スパース」とは、信号が殆どの時刻tにおいて0であることを指し、このスパース性は、例えば音声信号で確認される。

【0004】

なお、畳み込み混合の問題は扱いが複雑であること、また、スパース性の仮定は時間一周波数領域でよりよく成立することから、上述の式（1）に短時間フーリエ変換（DFT: Discrete Fourier Transform）を施し、信号を時間一周波数領域に変換した上で問題を扱うことが有効である。ここで、この混合過程を時間一周波数領域でモデル化すると、上述の式（1）は、

$$X(\omega, m) = H(\omega) S(\omega, m) \cdots (2)$$

となる。ここで、 ω は周波数、mはDFTに用いるフレーム時刻を表す。また、 $H(\omega)$ は、そのi行j列要素に信号源iからセンサjまでの周波数応答 $h_{ji}(\omega)$ を持つ（ $M \times N$ ）行列であり、以後これを混合行列と呼ぶ。また、 $S(\omega, m)$ 、 $X(\omega, m)$ は、それぞれ原信号、観測信号のDFT結果であり、

$$S(\omega, m) = [S_1(\omega, m), \dots, S_N(\omega, m)]^T$$

$$X(\omega, m) = [X_1(\omega, m), \dots, X_M(\omega, m)]^T$$

と表される。なお、記号 $[a]^T$ は a の転置行列を表す。

【0005】

〔分離過程のモデル化〕

次に、分離過程についてモデル化する。

ブラインド信号分離の目的は、原信号 s_i や混合過程 h_{ji} を知らずに、観測信号 x_j のみから、分離システム W_{203} を推定し、分離信号 y_k ($k=1, \dots, N$)を得ることである(図8の(a))。ここで、この分離過程を時間一周波数領域でモデル化すると、

$$Y(\omega, m) = W(\omega, m)X(\omega, m) \cdots (3)$$

となる。ここで、 $W(\omega, m)$ は、その j 行 k 列要素にセンサ j の観測信号から分離信号 y_k までの周波数応答 $W_{jk}(\omega, m)$ を持つ($N \times M$)行列であり、以後これを分離行列と呼ぶ。また、 $Y(\omega, m)$ は、分離信号のDFT結果であり、

$$Y(\omega, m) = [Y_1(\omega, m), \dots, Y_N(\omega, m)]^T$$

と表される。そして、これを短時間逆フーリエ変換(IDFT: Inverse Discrete Fourier Transform)することで、原信号の推定結果である分離信号 y_k が得られる。なお、分離された分離信号 y_k の順序は、原信号 s_i の順序と必ずしも一致しない。すなわち、分離信号 y_k は、原信号 s_k を推定したものとは限らない。

【0006】

〔分離行列 $W(\omega, m)$ の推定〕

ブラインド信号分離では、観測信号のみから分離行列 $W(\omega, m)$ を推定する。そして、この分離行列 $W(\omega, m)$ の推定方法には、(a)独立成分分析による方法、(b)信号のスパース性を利用した方法、(c)スパース性により混合行列を推定する方法が知られている。以下、これら従来の分離行列 $W(\omega, m)$ の推定方法について説明する。なお、以下では、特に断りのない限り、時間一周波数領域で説明を行うものとする。

【0007】

＜独立成分分析による方法(従来法1)＞

独立成分分析(ICA: Independent Component Analysis)とは、式(1)のように線形混合された信号を、信号の統計的独立性に基づいて分離する技術である(例えば、特許文献1参照)。特に、時間・周波数領域の独立成分分析では、各周波数における分離行列 $W(\omega, m)$ は、出力信号 $Y(\omega, m)$ の各要素が互いに独立となるよう、学習則 $W = W + \Delta W$ により逐次的に学習される。ここで、 ΔW は、

$$\Delta W = \mu [I - \phi[Y] Y^T]$$

という学習則により求められる。ここで I は単位行列、 $\langle \cdot \rangle$ は時間平均、 ϕ はある非線形関数、 μ は更新係数を、それぞれ表す。なお、ICAで求められる分離システムは、線形システムである。 $N=M=2$ の場合について、このICAによる分離過程のブロック図を、図8の(b)に示す。

なお、ICAでは、信号源の数 N とセンサ数 M について $M \geq N$ である必要がある。

【0008】

＜信号のスパース性を利用した方法(従来法2)＞

この方法では、信号のスパース性と相互独立性を仮定することで分離行列 $W(\omega, m)$ の推定を行う(例えば、非特許文献1参照)。なお、この方法は、信号源の数 N とセンサ数 M が $M \leq N$ の関係にある場合にも適用可能である。

以下、この信号のスパース性を利用した分離行列 $W(\omega, m)$ の推定方法を説明する。

上述のように、この方法では信号のスパース性と相互独立性を仮定する。この場合、たとえ複数の信号が同時に存在していたとしても、サンプルレベルでは、各信号が同時刻に互いに重なりあって観測される確率は低いことを仮定できる。つまり、この仮定のもとでは、各時刻で観測される信号は、何れか1つの信号源の信号のみで構成されている可能性が高い。

【0009】

従って、それぞれの時刻で観測された信号を、何らかの方法でクラスタリングし、このクラスごとに各時刻の信号を抽出するような関数（以下「バイナリマスク」という。）を求めることができれば、これを分離システム $W(\omega, m)$ として、観測信号を分離することができる。

各時刻での信号源の推定には、以下の方法を用いるのが一般的である。すなわち、それぞれの信号源が空間的に離れて配置されているとすると、複数のセンサで観測される信号間に、各信号源とセンサの相対位置によって決まる位相差や振幅比が発生する。従って、各時刻における観測信号の位相差や振幅比をクラスタリングし、それぞれのクラスに属する時刻の信号を再構成することによって、各原信号を推定することが可能となる。

【0010】

図9は、このような分離行列 $W(\omega, m)$ の推定過程を例示したブロック図である。なお、この図における「（従来法2）」の過程がここでの説明に該当する。

まず、各センサでの観測信号 $X(\omega, m)$ は、位相差・振幅比計算部 251 に送られ、そこで、

【数2】

$$\text{位相差 } \psi(\omega, m) = \angle \frac{X_1(\omega, m)}{X_2(\omega, m)} \text{ 及び振幅比 } \alpha(\omega, m) = \frac{|X_1(\omega, m)|}{|X_2(\omega, m)|}$$

の少なくとも一方が計算される。なお、この例では、センサ1 ($j=1$) 及びセンサ2 ($j=2$) での観測結果を用いている。

【0011】

この結果は、位相差・振幅比分布図作成部 252 に送られ、そこでその分布が計算される。ここで、この分布は、信号源の数 N 個のピークを持つ分布となる。図10に、この分布を例示する。なお、この例は、3信号の混合信号 ($N=3$) をセンサ1 ($j=1$) 及びセンサ2 ($j=2$) で観測した場合の例であり、(a) は位相差又は振幅比のみを用いて分布を求めた例、(b) は位相差と振幅比とを用いて分布を求めた例である。

位相差・振幅比分布図作成部 252 は、この分布から N 個のピーク（以下「分布の平均値」と呼ぶ。）を求め、その分布の平均値を、便宜上、値の小さい方から $\theta_1, \theta_2, \dots, \theta_N$ とする（例えば、図10参照。）。なお、「 θ 」は、

【数3】

「 θ 」を意味する。

【0012】

この分布の平均値 $\theta_1, \theta_2, \dots, \theta_N$ は、バイナリマスク作成部 253 に送られ、そこで以下のようなバイナリマスク $M_k(\omega, m)$ が作成される（例えば、図10のバイナリマスク 300、311～313参照。）。

【数4】

$$M_k(\omega, m) = \begin{cases} 1 & \theta_k - \varepsilon \leq \psi(\omega, m) \text{ or / and } \alpha(\omega, m) \leq \theta_k + \varepsilon \\ 0 & \text{otherwise} \end{cases} \quad (k=1, \dots, N) \dots (4)$$

そして、このバイナリマスクを用い、 $Y_k(\omega, m) = M_k(\omega, m) X_j(\omega, m)$ によって、 k 番目の分離信号を得ることができ、全ての分離信号 $Y(\omega, m)$ が復元できる。ここで、 j は、バイナリマスクを作用させる観測信号の番号となり、この例の場合は1或いは2の何れの値であってもよい。

即ち、この例のスパース性による方法では、分離行列 $W(\omega, m)$ は、 $W_{kj}(\omega, m) = M_k(\omega, m)$ for $j=1$ or 2 、 $W_{kl}(\omega, m) = 0$ for $l \neq j$ ($l=1, \dots, M$) の非線型システムとなる。

【0013】

<スパース性により混合行列を推定する方法（従来法3）>

この方法では、信号のスパース性を用いて混合行列 $H(\omega)$ を推定し、その逆行列を用いて信号を分離する（例えば、非特許文献2参照。）。また、この方法は、推定した混合行

列 $H(\omega)$ の逆行列を用いるため、信号源の数 N とセンサ数 M が $M=N$ の関係にある場合のみ適用可能である。以下、この方法の説明を行う。

図9は、このような分離信号 $Y(\omega, m)$ の推定過程を例示したブロック図である。なお、この図における「(従来法3)」の過程がここでの説明に該当する。

【0014】

混合信号 $X(\omega, m)$ は、混合行列 $H(\omega)$ を用いて

【数5】

$$\begin{bmatrix} X_1(\omega, m) \\ X_2(\omega, m) \\ \vdots \\ X_N(\omega, m) \end{bmatrix} = \begin{bmatrix} H_{11}(\omega) & \cdots & H_{1N}(\omega) \\ H_{21}(\omega) & \cdots & H_{2N}(\omega) \\ \vdots & \ddots & \vdots \\ H_{N1}(\omega) & \cdots & H_{NN}(\omega) \end{bmatrix} \begin{bmatrix} S_1(\omega, m) \\ S_2(\omega, m) \\ \vdots \\ S_N(\omega, m) \end{bmatrix} \cdots (5)$$

$$= \begin{bmatrix} 1 & \cdots & 1 \\ H_{21}(\omega)/H_{11}(\omega) & \cdots & H_{2N}(\omega)/H_{1N}(\omega) \\ \vdots & \ddots & \vdots \\ H_{N1}(\omega)/H_{11}(\omega) & \cdots & H_{NN}(\omega)/H_{1N}(\omega) \end{bmatrix} \begin{bmatrix} H_{11}(\omega)S_1(\omega, m) \\ H_{12}(\omega)S_2(\omega, m) \\ \vdots \\ H_{1N}(\omega)S_N(\omega, m) \end{bmatrix} \quad (6)$$

$$\equiv \hat{H}(\omega) \hat{S}(\omega, m) \cdots (7)$$

と表される。そのため、 $\hat{H}(\omega)$ を推定できれば、

$$Y(\omega, m) = \hat{S}(\omega, m) = \hat{H}(\omega)^{-1} X(\omega, m) \cdots (8)$$

により、分離信号 $Y(\omega, m)$ を推定できる。以下、この $\hat{H}(\omega)$ の推定から分離信号 $Y(\omega,$

$m)$ までの流れを説明する。なお、以下において、 α^\wedge は、 α を示す。

【0015】

はじめに、位相差・振幅比計算部251、位相差・振幅比分布図作成部252及びバイナリマスク作成部253を経て、上述の式(4)で表されるバイナリマスク $M_k(\omega, m)$ を生成する。そして、このバイナリマスク $M_k(\omega, m)$ と観測信号 $X(\omega, m)$ を用い、1つの信号源のみから信号が発せられている時刻の信号を求めると、

【数6】

$$\hat{X}(\omega, m) = M_k(\omega, m) X(\omega, m) \cdots (9)$$

となる。なお、ここでは、すべてのセンサの観測信号 $X(\omega, m) = [X_1(\omega, m), \cdots, X_M(\omega, m)]^T$ についてバイナリマスク $M_k(\omega, m)$ を適用する。ここで、例えば、1つの原信号 $S_i(\omega, m)$ のみが発せられている時刻 m_i の信号を、観測信号 $X_j(\omega, m_i)$ を用いて算出した結果は、

【数 7】

$$\hat{X}_j(\omega, m_i) = M_i(\omega, m_i) X_j(\omega, m_i) \cdots (10)$$

となる。

【0016】

そして、 S_i 以外の原信号が 0 であること、及び上記の式 (5) から、

【数 8】

$$\hat{X}_j(\omega, m_i) = H_{ji}(\omega) S_i(\omega, m_i) \cdots (11)$$

が成立し、式 (10) (11) より、

【数 9】

$$\hat{X}_j(\omega, m_i) = M_i(\omega, m_i) X_j(\omega, m_i) = H_{ji}(\omega) S_i(\omega, m_i) \cdots (12)$$

が成立する。

このように求められた分離信号は混合過程計算部 254 に送られ、そこで

$$\hat{H}_{ji}(\omega) = E \left[\frac{\hat{X}_j(\omega, m_i)}{\hat{X}_i(\omega, m_i)} \right] = E \left[\frac{M_i(\omega, m_i) X_j(\omega, m_i)}{M_i(\omega, m_i) X_i(\omega, m_i)} \right] \cdots (13)$$

の演算が行われ、 $\hat{H}(\omega)$ が算出される。ここで、 $E[\cdot]$ は、 m_i に関する平均である。

なお、この式 (13) は、上述の式 (6)、(7) より導出される

$$\hat{H}_{ji}(\omega) = E \left[\frac{H_{ji}(\omega)}{H_{ii}(\omega)} \right] = E \left[\frac{H_{ji}(\omega) S_i(\omega, m_i)}{H_{ii}(\omega) S_i(\omega, m_i)} \right] \cdots (14)$$

に、上記の式 (12) を代入することにより導き出されるものである。

このように求められた $\hat{H}(\omega)$ は、逆行列計算部 255 に送られ、そこでその逆行列

$\hat{H}(\omega)^{-1}$ が求められる。そして、この逆行列を上記の式 (8) に代入することにより、分離信号 $Y(\omega, m)$ の推定ができる。

【特許文献 1】特開 2003-84793 号公報

【非特許文献 1】S. Rickard and O. Yilmaz, "On the approximate W-Disjoint orthogonality of speech," in Proc., ICASSP2002, 2002, vol. 1, pp. 529-532

【非特許文献 2】Y. Deville, "Temporal and time-frequency correlation-based blind source separation methods," in Proc., ICASSP2003, Apr. 2003, pp. 1059

-1064

【発明の開示】

【発明が解決しようとする課題】

【0017】

しかし、従来の分離行列の推定方法では、信号源の数 N とセンサの数 M が $N > M$ の関係にある場合に、混合信号を高い品質で分離することは困難であるという問題がある。

すなわち、前述のように、信号源の数 N とセンサの数 M が $N > M$ の関係にある場合、独立成分分析による方法、及びスパース性により混合行列を推定する方法は使用できない。また、信号のスパース性を利用した方法は使用できるが、この方法では、分離性能がよく、なおかつ歪みが小さい信号分離を行うことが困難である。

【0018】

つまり、例えば、上述の式(4)で示されるバイナリマスクの作成時、 ϵ を十分小さくすると良い分離性能を得ることができるが、その反面、このバイナリマスクによって取り出される信号サンプルが少なくなり、本来原信号の成分として取り出されるべき成分の多くが取り出されない。この場合、各分離信号には多くの0成分が詰めこまれることになり、これが原因となって、歪みが大きくなり、Musical Noiseと呼ばれる聴感上不快なノイズが発生する。一方、この歪みを回避するためバイナリマスクの ϵ を大きくすると、分離信号に詰められる0が少なくなるため、Musical Noiseは減少するが、分離性能は劣化する。

【0019】

この発明はこのように鑑みてなされたものであり、信号源の数 N とセンサの数 M が $N > M$ の関係にある場合に、混合信号を高い品質で分離することを可能にするブラインド信号分離装置を提供することを目的とする。

また、この発明の他の目的は、信号源の数 N とセンサの数 M が $N > M$ の関係にある場合に、混合信号を高い品質で分離することを可能にするブラインド信号分離方法を提供することである。

さらに、この発明の他の目的は、信号源の数 N とセンサの数 M が $N > M$ の関係にある場合に、混合信号を高い品質で分離することを可能にする機能をコンピュータに実行させるためのブラインド信号分離プログラムを提供することである。

【課題を解決するための手段】

【0020】

この発明では上記課題を解決するために、まず、センサで観測した観測信号を、周波数領域の信号に変換し、この周波数領域の信号を用い、センサ間における観測値の相対値を、各周波数について算出する。次に、相対値の分布から N 個のピーク位置（以下「平均値」という。）を算出し、この平均値を用い、変換した周波数領域の信号から、2個以上 M 個以下の信号源から発せられた信号によって構成される混合信号（以下「限定混合信号」という。）を抽出するためのバイナリマスクを作成する。そして、この作成したバイナリマスクを用い、上述の周波数領域の信号から、限定混合信号を抽出し、この限定混合信号をそれぞれの信号に分離して分離信号を得た後、この分離信号を時間領域の信号に戻す。

【0021】

ここで、混合信号を抽出するためのバイナリマスクを作成し、これによって限定混合信号を抽出することにより、この限定混合信号（2個以上 M 個以下の信号源から発せられた信号によって構成）に対し、信号源の数 N とセンサの数 M が $N \leq M$ の関係にある場合にしか使用できない分離行列の推定方法を適用することで、 $N > M$ の場合の信号分離が可能となる。

また、この発明において、1つのみのバイナリマスクを用いた場合、 N 個のうち M 個以下の信号しか分離できない。そのため、好ましくは、複数通りのバイナリマスクを作成し、これら複数のバイナリマスクを用い、複数種類の限定混合信号を抽出する。これにより、全ての原信号を何れかの限定混合信号に混合させることが可能となり、全ての原信号を分離推定することが可能となる。

【発明の効果】

【0022】

この発明では、2個以上M個以下の信号源から発せられた信号によって構成される限定混合信号を抽出するためのバイナリマスクを作成し、これによって限定混合信号を抽出する。そのため、この限定混合信号に、信号源の数Nとセンサの数Mが $N \leq M$ の関係にある場合にしか使用できない分離行列の推定方法を適用することが可能となる。その結果、 $N > M$ の場合の分離が可能となる。

【0023】

また、限定混合信号は、2個以上の信号源から発せられた信号によって構成される。そのため、この限定混合信号を抽出するためのバイナリマスクで取り出した信号は、1つの原信号のみを抽出するための従来のバイナリマスクで取り出した信号に比べ、歪みが小さい（詳細は後述する。）。

以上より、信号源の数Nとセンサの数Mが $N > M$ の関係にある場合でも、混合信号を高い品質で分離することが可能となる。

【発明を実施するための最良の形態】

【0024】

以下、この発明の実施の形態を図面を参照して説明する。

〔原理〕

まず、この形態の原理について説明する。

この形態は、N個の信号源から発せられた信号が、M ($M < N$) 個のセンサで混合して観測される状況で適用されるものである。そして、スパース性に基づくバイナリマスクを、個々の目的信号（分離信号）抽出のためではなく、2個以上M個以下の原信号からなる混合信号（前述の「限定混合信号」に相当）の抽出のために用いる。具体的には、例えば、バイナリマスクを、N個の信号が混合された観測信号から $N - M$ 個の信号を除去するために用い、N個の信号が混合された観測信号から $N - M$ 個の信号を除去することにより、M個の原信号からなる限定混合信号を得る。

【0025】

このような限定混合信号は、M個以下の原信号によって構成されるため、「独立成分分析による方法」や「スパース性により混合行列を推定する方法」を適用できる。これらは、原理的にMusical Noiseが起らない手法であり、これらを用いることにより、信号源の数Nとセンサの数Mが $N > M$ の関係にある場合でも、混合信号を高い品質で分離することができる。

しかし、この場合、限定混合信号にはM個以下の原信号成分しか含まれておらず、これだけではM個以下の分離信号しか得られない。そのため、N個全ての分離信号が必要な場合は、さらにバイナリマスクの作り方を変えて、限定混合信号として取り出す原信号の構成を変化させる。そして、この再構成した限定混合信号に対し、「独立成分分析による方法」等を適用して分離行列を再推定し、他の原信号に対応する分離信号を抽出する。

【0026】

〔歪が小さくなる仕組〕

以下に、この形態を適用した分離信号の歪が小さくなる仕組を説明する。

前述のように、従来の「信号のスパース性を利用した方法」で分離性能を上げた際に分離信号の歪が大きくなるのは、分離性能を上げるために上述の式(4) ϵ を十分小さくすると、このバイナリマスクによって取り出される信号成分が制限され、本来原信号の成分として取り出されるべき成分の多くが取り出されないことになるからである。つまり、この場合、各分離信号に多くの0成分が詰めこまれ、各分離信号を不連続とし、Musical Noiseを発生させることになるからである。

これに対し、この形態では、バイナリマスクによって、2個以上M個以下の原信号からなる混合信号（限定混合信号）を抽出する。具体的には、例えば、バイナリマスクを用い、観測信号から除去する信号を抽出し、観測信号からこの抽出信号を除去することにより限定混合信号を得る手法（手法1）、或いは、バイナリマスクを用い、観測信号から直接

、2個以上の原信号からなる限定混合信号を抽出する手法（手法2）等が採られる。

【0027】

手法1の場合、バイナリマスクによって取り出されるのは「観測信号から除去する信号」であるため、このバイナリマスクによって取り出される信号幅を狭くし、取り出されるべき成分の一部が取り出されないこととなっても、最終的に分離する信号の成分が排除されるわけではない。よって分離信号に不適切な0が詰め込まれることは少なく、Musical Noiseの発生も少ない。

また、手法2の場合、限定混合信号を抽出するバイナリマスクは、従来の「信号のスパース性を利用した方法（従来法2）」に比べて分布中の広い範囲の信号を抽出するので、バイナリマスクにより取り出される信号は従来法2に比べ多い。よって、手法2の場合には、分離信号に不適切に詰め込まれる0が少なくなるため、Musical Noiseも小さくなる。

【0028】

〔バイナリマスクによる0詰めの影響の検証〕

以下に、3人の話者による音声信号 s_1 、 s_2 、 s_3 が、2つの無指向性マイクで観測された場合（ $N=3$ 、 $M=2$ の場合）について、バイナリマスクによる0詰めの影響を議論する。

バイナリマスクによって0が詰められ失われた信号のパワーの比率を、

【数10】

$$\frac{\sum_t |s_i(t)|^2 - \sum_t |y_i(t)|^2}{\sum_t |s_i(t)|^2} \times 100 \dots (15)$$

と定義すると、従来の「信号のスパース性を利用した方法（従来法2）」では、 $s_1: 17\%$ 、 $s_2: 14\%$ 、 $s_3: 23\%$ もの信号のパワーがバイナリマスクにより失われた。

【0029】

一方、この例の発明法のバイナリマスクによる信号のパワー劣化は、限定混合信号を s_1 と s_2 の2つの信号の混合とした場合には $s_1: 2.5\%$ 、 $s_2: 5.7\%$ であり、 s_2 と s_3 の2つの信号の混合とした場合には $s_2: 8.1\%$ 、 $s_3: 0.7\%$ であった。

すなわち、この形態では、バイナリマスクによる信号の劣化が従来法に比べて少ないことが分かる。これは、この形態では、Musical Noiseが発生しにくいことを示している。

【0030】

次に、この実施形態について説明する。

〔第1の実施の形態〕

図1は、公知のコンピュータに所定のプログラムを実行させることにより構成される、この形態のブラインド信号分離装置1の全体構成を例示したブロック図であり、図2は、このブラインド信号分離装置1を構成する信号選択部20及び信号分離部30の詳細構成を例示したブロック図である。また、図3は、この形態におけるブラインド信号分離方法説明するためのフローチャートの例示である。以下、これらの図を用い、この形態の機能構成及びその処理について説明を行っていく。

この例のブラインド信号分離装置1は、 N 個の信号源2から発せられた原信号 $s_i(t)$ ($i=1, \dots, N$) が、 M ($M < N$) 個のセンサ3で混合して観測される状況において、その観測結果から、制御部50の制御の下、所定の信号を分離する。

【0031】

各センサ3で観測された時間領域の観測信号 $x_j(t)$ ($j=1, \dots, M$) は、ブラインド信号分離装置1の周波数領域変換部10に入力され（ステップS1）、そこで例えば、短時間フーリエ変換によって、周波数領域の信号 $X_j(\omega, m)$ ($j=1, \dots, M$) に変換される（ステップS2）。

周波数領域変換部10で変換された周波数領域の信号 $X_j(\omega, m)$ は、信号選択部20の相対値算出部21に送られ、そこで、この周波数領域の信号 $X_j(\omega, m)$ を用い、

各センサ3間における観測値の相対値を、各周波数について算出する(ステップS3)。

【0032】

ここで、この観測値の相対値とは、例えば、各センサ3の観測信号間の

【数11】

$$\text{位相差 } \psi(\omega, m) = \angle \frac{X_{j_1}(\omega, m)}{X_{j_2}(\omega, m)} \text{ 及び 振幅比 } \alpha(\omega, m) = \frac{|X_{j_1}(\omega, m)|}{|X_{j_2}(\omega, m)|}$$

の少なくとも一方である(j_1 及び j_2 は、対比されるセンサ3の番号)。例えば、無指向性のセンサ3を用いる場合は、この相対値として位相差 $\psi(\omega, m)$ を用いることが好ましく、指向性のセンサ3を用いる場合は、この相対値として振幅比 $\alpha(\omega, m)$ を用いるか、位相差 $\psi(\omega, m)$ と振幅比 $\alpha(\omega, m)$ の両方を用いることが好ましい。

このように相対値算出部21で算出された位相差 $\psi(\omega, m)$ 、振幅比 $\alpha(\omega, m)$ 等の相対値の情報は、平均値算出部22に送られ、そこで、得た観測信号フレーム数分のデータを用い、これら相対値の分布図(ヒストグラム)が作成される(ステップS4)。図4は、このように平均値算出部22で作成された相対値の分布図の例示である。図4の例では、位相差 $\psi(\omega, m)$ 或いは振幅比 $\alpha(\omega, m)$ の何れか一方のみを「相対値」として分布図を作成しており、横軸に位相差 $\psi(\omega, m)$ 或いは振幅比 $\alpha(\omega, m)$ を、縦軸にカウント数をとっている。なお、この図は、信号源2の数 $N=3$ 、センサ3の数 $M=2$ の場合の例である。

【0033】

次に平均値算出部22は、作成したヒストグラムのピークを、 N 個の分布の平均値の近似値とみなし、この平均値を算出する(ステップS5)。図4の例では、相対値 θ_1 、 θ_2 及び θ_3 の3箇所に分布のピークがあり、この相対値 θ_1 、 θ_2 及び θ_3 が「平均値」となる。なお、平均値 θ_u ($u=1, \dots, N$)は、値が小さい方から順(昇順)に、 $\theta_1, \dots, \theta_N$ と表す。これにより、後述する複数のバイナリマスクの構成が容易となる。

このように算出された N 個の平均値 $[\theta_1, \dots, \theta_N]$ は、バイナリマスク作成部23に送られ、この例のバイナリマスク作成部23は、この平均値 $[\theta_1, \dots, \theta_N]$ を用い、周波数領域変換部10によって変換された周波数領域の信号 $X_j(\omega, m)$ ($j=1, \dots, M$)から、2個以上 M 個以下の信号源2から発せられた信号によって構成される混合信号(限定混合信号)を抽出するための、バイナリマスク $B(\omega, m)$ を作成する(ステップS6)。以下、この形態のバイナリマスクについて説明する。

【0034】

<この形態のバイナリマスク>

この形態で作成されるバイナリマスク $B(\omega, m)$ は、所定の範囲内(θ_{\min} 以上 θ_{\max} 以下:「第1の平均値と第2の平均値を含む所定の範囲内」に相当)にある相対値をとる(ω, m)では0以外の数値(例えば1)を出力し、この範囲内にない相対値をとる(ω, m)では0を出力する関数である。すなわち、例えば、

【数12】

$$B(\omega, m) = \begin{cases} 1 & \theta_{\min} \leq \psi(\omega, m) \text{ or/and } \alpha(\omega, m) \leq \theta_{\max} \\ 0 & \text{otherwise} \end{cases} \dots (16)$$

をバイナリマスクとして作成する(ここで「 $\theta_{\min} \leq \psi(\omega, m) \text{ or/and } \alpha(\omega, m) \leq \theta_{\max}$ 」とは、 $\psi(\omega, m)$ 及び $\alpha(\omega, m)$ の少なくとも一方が、 θ_{\min} 以上 θ_{\max} 以下の範囲に属することを意味する。)。ここで θ_{\min} 、 θ_{\max} は、例えば、 $k=k_1, \dots, k_M$ を抽出する場合、 $\theta_{k_1-1} \leq \theta_{\min} \leq \theta_{k_1}$ 、 $\theta_M \leq \theta_{\max} \leq \theta_{M+1}$ となる範囲で設定する(添え字の「 k_1-1 」は、 k_1-1 を示す。)。なお、この例では、 θ_{k_1} が「第1の平均値」に相当し、 θ_M が「第2の平均値」に相当する。

【0035】

また、 θ_{\min} から θ_{\max} の範囲に含まれる平均値（例えば、「第1の平均値から第2の平均値までの範囲に含まれる平均値」）の数 Z は、センサの数 M 以下であり、好ましくは、センサの数 M である。例えば、図4の(a)に示した $M=2$ の場合のバイナリマスク61は、平均値 θ_1 及び平均値 θ_2 の2つの平均値を含んでいる。

さらに、この例のバイナリマスク作成部23は、複数通りのバイナリマスク $B(\omega, m)$ を作成する。具体的には、この例のバイナリマスク作成部23は、上述したバイナリマスクに含まれる平均値の組み合わせを変更（「第1の平均値及び第2の平均値となる平均値の組み合わせを変更」に相当。）することにより、複数通りのバイナリマスクを作成する。また、好ましくは、この例のバイナリマスク作成部23は、第1の平均値と第2の平均値との間に含まれない平均値を含むよう、新たな第1の平均値或いは第2の平均値を設定して選択していくことにより、複数通りの上記バイナリマスクを作成する。

【0036】

図4の例を用いて説明すると、図4の(a)に例示したバイナリマスク61は、 $k=1, 2$ の場合について $\theta_{\min}, \theta_{\max}$ を設定し、上述の式(16)により作成したバイナリマスクである。一方、図4の(b)に例示したバイナリマスク62は、 θ_{\min} と θ_{\max} との範囲に含まれない平均値 θ_3 を含むよう、 $k=2, 3$ として $\theta_{\min}, \theta_{\max}$ を設定し、上述の式(16)により作成したバイナリマスクである。

この例のバイナリマスク作成部23は、このようにして複数種類のバイナリマスク $B(\omega, m)$ を作成していき、作成したバイナリマスク $B(\omega, m)$ を、例えばバイナリマスク作成部23のメモリに記録していく（この形態のバイナリマスクの説明終わり。）。

【0037】

この例のバイナリマスク作成部23で作成されたバイナリマスク $B(\omega, m)$ の1つは、限定混合信号抽出部24に送られ、そこで、このバイナリマスク $B(\omega, m)$ を用い、周波数領域の観測信号 $X_j(\omega, m)$ ($j=1, \dots, M$) から、上述の限定混合信号を抽出する（ステップS7）。この例の限定混合信号抽出部24は、周波数領域の観測信号 $X_j(\omega, m)$ にバイナリマスク $B(\omega, m)$ を乗じた値を、限定混合信号 $X^*(\omega, m)=[X^*_1(\omega, m), \dots, X^*_M(\omega, m)]^T$ と近似して抽出する（ $X^*(\omega, m)=B(\omega, m) X(\omega, m)$ ）。なお、図4の例においてバイナリマスク61を適用した場合には、平均値 θ_1, θ_2 に対応する原信号で構成される限定混合信号が抽出され、バイナリマスク62を適用した場合には、平均値 θ_2, θ_3 に対応する原信号で構成される限定混合信号が抽出されることになる。

【0038】

このようにして限定混合信号抽出部24で抽出された限定混合信号 $X^*(\omega, m)$ は、信号分離部30に送られ、この例の信号分離部30は、ICAを用いて、この限定混合信号 $X^*(\omega, m)$ の分離信号 $Y(\omega, m)=[Y_1(\omega, m), \dots, Y_Z(\omega, m)]^T$ を算出する（ステップS8）。つまり、信号分離部30は、分離信号 $Y(\omega, m)$ の各要素が互いに独立となるように、学習則 $W=W+\Delta W$ を逐次適用して分離行列 $W(\omega, m)$ を決定し、分離信号 $Y(\omega, m)$ を算出する（例えば、前述の特許文献1参照。）。図2を用いて説明すると、例えばまず、この限定混合信号 $X^*(\omega, m)$ は、分離行列計算部31に送られ、分離行列計算部31は、分離行列 $W(\omega, m)$ によって限定混合信号 $X^*(\omega, m)$ を分離する。この分離信号は、活性化関数計算部32に送られ、活性化関数計算部32は、この分離信号に非線形の活性化関数 ϕ を作用させる。この演算結果は、独立性判定部33に送られ、独立性判定部33は、自然勾配法と呼ばれる $\Delta W=\mu[I-\langle\phi[Y]Y^T\rangle]$ の式に従い、 ΔW を算出する（ μ は更新係数、 I は単位行列、 $\langle\alpha\rangle$ は α の時間平均を、それぞれ表す。）。算出された ΔW は分離行列計算部31に送られ、分離行列計算部31は、学習則 $W=W+\Delta W$ に従って分離行列 $W(\omega, m)$ を改良する。そして、以後、上述と同様な処理が所定回繰り返され、分離行列計算部31から分離信号 $Y(\omega, m)$ が出力される。なお、 Z は、2以上 M 以下の自然数であり、限定混合信号に含まれると推定される原信号の数を示す。

分離行列計算部31から出力された分離信号 $Y(\omega, m)=[Y_1(\omega, m), \dots, Y_Z(\omega, m)]^T$ は、時間領域変換部40に送られ、そこで、例えば、短時間逆フーリエ変換等によって、時間

領域の分離信号 $y_1(t), \dots, y_z(t)$ に変換されて (ステップ S9) 出力される (ステップ S10)。

【0039】

次に、制御部 50 において、必要な分離信号が全て抽出済みであるか否かを判断する (ステップ S11)。つまり、ステップ S10 で出力される分離信号 $y_1(t), \dots, y_z(t)$ は、限定混合信号 $X^*(\omega, m)$ の分離信号であり、限定混合信号 $X^*(\omega, m)$ に含まれない原信号は出力されない。よって、制御部 50 で、必要な原信号に対応する分離信号が全て抽出されたか否かを判断し、ここで、必要な分離信号が全て抽出済みであれば処理を終了し、そうでなければ、バイナリマスク作成部 23 のメモリに格納されている他のバイナリマスク $B(\omega, m)$ を抽出し、バイナリマスク $B(\omega, m)$ を変換した後に (ステップ S12)、ステップ S7 の処理に戻る。

【0040】

以下に、この形態のシミュレーション結果を示す。

【表1】

	SIR1	SIR2	SIR3	SDR1	SDR2	SDR3
従来法	15.4	10.3	14.6	9.8	11.9	9.2
本形態	—	8.4	16.4	—	15.0	20.9
	13.1	8.2	—	17.4	13.8	—

この例では、原信号として、3 人の話者 (男性 2 名・女性 1 名) による音声信号を用い、残響の無い環境でのこれらの混合信号を、2 つの無指向性マイクで観測する状況をシミュレートしている。なお、この表の SIR は信号対妨害音比 (Signal to interference ratio) (dB) であり、分離性能を示す指標である。また、SDR は信号対歪比 (Signal to distortion ratio) (dB) であり、信号の歪の程度を示す指標である。双方とも値が大きい方が性能が良いことを示している。また、SIR1 及び SDR1 は話者 1 に、SIR2 及び SDR2 は話者 2 に、SIR3 及び SDR3 は話者 3 に、それぞれ対応している。この表に示すように、この形態の方法では、分離性能 SIR をほとんど落すことなく、従来法 2 に比べ格段に高い SDR を得ることができる。これは、この形態の方法が、格段に低い歪で信号の分離を行っていることを示している。

【0041】

〔第 2 の実施の形態〕

この形態は、第 1 の実施の形態の変形例であり、バイナリマスク作成部 23 で 2 種類のバイナリマスクを作成し、信号分離部 30 で「スパース性により混合行列を推定する方法 (例えば、非特許文献 2 参照。)」を適用する点が第 1 の実施の形態と相違する。以下では、第 1 の実施の形態との相違点を中心に説明を行い、第 1 の実施の形態と共通する事項については説明を省略する。なお、第 1 の実施の形態と機能構成が共通する部分は、ブロック図において第 1 の実施の形態のものと同一の符号を付した。

図 5 は、この形態におけるブラインド信号分離装置 1 の信号選択部 20 と信号分離部 30 の構成を例示したブロック図であり、図 6 は、この形態におけるブラインド信号分離方法を説明するためのフローチャートの例示である。以下、これらの図を用いて、この形態の機能構成及びその処理について説明を行っていく。

【0042】

まず、第 1 の実施の形態と同様に、各センサ 3 で観測された時間領域の観測信号 $x_j(t)$ が、周波数領域変換部 10 に入力され (ステップ S21)、周波数領域の信号 $X_j(\omega, m)$ ($j=1, \dots, M$) に変換される (ステップ S22)。そして、相対値算出部 21 で、この信号 $X_j(\omega, m)$ を用いて各センサ 3 間における観測値の相対値を各周波数について算出し (ステップ S23)、平均値算出部 22 で、これら相対値の分布図が作成され (ステップ S24)、さらに平均値が算出される (ステップ S25)。

次に、バイナリマスク作成部 23 で 1 種類目のバイナリマスク $B(\omega, m)$ を作成する (ステップ S26)。このバイナリマスク $B(\omega, m)$ は、限定混合信号を抽出するバイ

ナリマスクであり、例えば、第1の実施の形態と同様、上述の式(16)を用いて設計する。ただし、この形態で作成されるバイナリマスク $B(\omega, m)$ は、 M 個の原信号のみからなる限定混合信号を抽出するためのバイナリマスクに限られる。すなわち、 θ_{\min} から θ_{\max} の範囲に含まれる平均値の数は、センサの数 M に限られる。なお、このバイナリマスク $B(\omega, m)$ は、第1の実施の形態と同様、複数作成される。

また、バイナリマスク作成部23は、2種類目のバイナリマスク $M_k(\omega, m)$ も作成する(ステップS27)。このバイナリマスク $M_k(\omega, m)$ は、1つの原信号のみを含む信号を抽出するバイナリマスクであり、例えば、前述の式(4)を用いて設計される。

【0043】

次に、信号分離部30において、例えば第1の実施の形態と同様に、観測信号 $X_j(\omega, m)$ にバイナリマスク $B(\omega, m)$ を適用して限定混合信号を取り出す(ステップS28)。なお、ここで取り出される限定混合信号は、 M 個の信号源から発せられた信号によって構成される混合信号になる。

このようにして限定混合信号抽出部24で抽出された限定混合信号 $\hat{X}(\omega, m)$ は、信号分離部30に送られ、この例の信号分離部30は、前述の「スパース性により混合行列を推定する方法」を用いて、この限定混合信号 $\hat{X}(\omega, m)$ の分離信号 $Y(\omega, m)=[Y_1(\omega, m), \dots, Y_z(\omega, m)]^T$ を算出する(ステップS29)。

【0044】

すなわち、信号のスパース性を用いて混合行列 $H(\omega)$ を推定し、その逆行列を用いて限定混合信号 $\hat{X}(\omega, m)$ を分離する。具体的には、例えば、まず、積演算部131において、バイナリマスク作成部22で作成されたバイナリマスク $M_k(\omega, m)$ と観測信号 $X(\omega, m)$ を用い、 $M_k(\omega, m)X(\omega, m)$ を計算する。この演算結果は混合仮定計算部132に送られ、そこで

【数13】

$$\hat{H}_{ji}(\omega) = E \left[\frac{M_i(\omega, m_i)X_j(\omega, m_i)}{M_i(\omega, m_i)X_1(\omega, m_i)} \right] \cdot \dots \quad (17)$$

の演算が行われ、 $\hat{H}(\omega)$ が算出される。ここで、 $E[\cdot]$ は、 m_i に関する平均である。

【0045】

このように求められた $\hat{H}(\omega)$ は、逆行列計算部133に送られ、そこでその逆行列が求められる。ただし、混合仮定計算部132で求められた $\hat{H}(\omega)$ は、 $N \times M$ のサイズの行列である。よって、逆行列計算部133は、逆行列の計算が可能のように、この $\hat{H}(\omega)$ のランクを落とす。すなわち、 $\hat{H}(\omega)$ の k 列目($k=k_1, \dots, k_M$)のみを取り出すことにより、 $M \times M$ の行列 $\hat{H}_M(\omega)$ を作成する。そして、逆行列計算部133は、この逆行列 $\hat{H}_M(\omega)^{-1}$ を作成し、積演算部134に送る。

積演算部134は、この逆行列 $\hat{H}_M(\omega)^{-1}$ と、信号選択部20から受け取った限定混合信号 $\hat{X}(\omega, m)$ とを用い、

$$Y(\omega, m) = \hat{H}_M(\omega)^{-1} \hat{X}(\omega, m)$$

の演算を行い、 M 個の分離信号 $Y(\omega, m)$ を推定する。

【0046】

積演算部134から出力された分離信号 $Y(\omega, m)$ は、時間領域変換部40に送られ、そこで、例えば、短時間逆フーリエ変換等によって、時間領域の分離信号 $y_1(t), \dots, y_M(t)$ に変換されて出力される(ステップS31)。

次に、第1の実施の形態と同様、制御部50において、必要な分離信号が全て抽出済みであるか否かを判断し(ステップS32)、ここで、必要な分離信号が全て抽出済みであれば処理を終了し、そうでなければ、バイナリマスク作成部23のメモリに格納されている他のバイナリマスク $B(\omega, m)$ を抽出し、バイナリマスク $B(\omega, m)$ を変換した後に

(ステップS33)、ステップS28の処理に戻る。

【0047】

〔第3の実施の形態〕

この形態は、上述の第1の実施の形態或いは第2の実施の形態の変形例であり、バイナリマスク作成部23が作成するバイナリマスク $B(\omega, m)$ の作成方法に違いがある。以下では、このバイナリマスク $B(\omega, m)$ の作成方法、及び限定混合信号の算出処理にのみについて説明を行い、その他の処理及び機能構成については、第1の実施の形態或いは第2の実施の形態と同一であるため、説明を省略する。

この形態のバイナリマスク $B(\omega, m)$ は、上述の限定混合信号以外の観測信号成分を抽出するためのものである。すなわち、この形態の限定混合信号抽出部24は、周波数領域変換10において変換された周波数領域の信号から、この周波数領域の信号にバイナリマスク $B(\omega, m)$ を乗じた値を減じることにより、限定混合信号を抽出する。例えば、上述の式(4)で示されるバイナリマスク $M_k(\omega, m)$ を $k=1, \dots, N-M$ まで作成し、

【数14】

$$\hat{X}(\omega, m) = X(\omega, m) - \sum_{k=1}^{N-M} \{M_k(\omega, m)X(\omega, m)\} \cdots (18)$$

を計算することで、M個の原信号のみからなる限定混合信号 $\hat{X}(\omega, m)$ を推定する。

【0048】

図7は、このバイナリマスク161、162を例示した図である。

この例は、 $N=4$ 、 $M=2$ の場合の例示であり、バイナリマスク161がバイナリマスク $M_1(\omega, m)$ に、バイナリマスク162がバイナリマスク $M_2(\omega, m)$ に、それぞれ該当する。このバイナリマスク161、162を用いて式(18)を計算することにより、平均値 θ_1 、 θ_2 に対応する2つの原信号からなる限定混合信号 $\hat{X}(\omega, m)$ を推定することができる。

限定混合信号 $\hat{X}(\omega, m)$ が算出されると、以後第1の実施の形態或いは第2の実施の形態と同様な処理が行われる。

【0049】

以上説明したように、上述の形態では、周波数領域変換部10において、観測信号を周波数領域の信号に変換し、相対値算出部21において、この周波数領域の信号を用い、センサ3間における観測値の相対値を、各周波数について算出する。そして、平均値算出部22において上述の平均値を算出し、バイナリマスク作成部23で、この平均値を用いて限定混合信号を抽出するためのバイナリマスクを作成し、限定混合信号抽出部24において、このバイナリマスクを用いて限定混合信号を抽出することとした。そのため、この限定混合信号に「独立成分分析による方法」等の原理的にMusical Noiseが起こらない手法を適用することが可能となる。その結果、信号源の数Nとセンサの数Mが $N > M$ の関係にある場合でも、混合信号を高い品質で分離することができる。

また、限定混合信号を、M個の信号源から発せられた信号によって構成される混合信号とすることにより、この限定混合信号に「スパース性により混合行列を推定する方法」も適用することも可能となる。

【0050】

さらに、上述の形態の例では、バイナリマスクを、第1の平均値と第2の平均値を含む所定の範囲内にある相対値に対して0以外の数値を出力し、この範囲内でない相対値に対して0を出力する関数とした。そして、この際、第1の平均値から第2の平均値までの範囲に含まれる平均値の数をセンサ3の数M以下とし、周波数領域の観測信号にバイナリマスクを乗じた値を、限定混合信号とすることにより、M個以下の原信号が混合した限定混合信号を得ることができる。これにより、この限定混合信号に「独立成分分析による方法」等の原理的にMusical Noiseが起こらない手法を適用することが可能となる。その結果、信号源の数Nとセンサの数Mが $N > M$ の関係にある場合でも、混合信号を高い品質で分離することができる。

さらに、第1の平均値から第2の平均値までの範囲に含まれる平均値数をセンサ3の数Mとすることにより、この限定混合信号に「スパース性により混合行列を推定する方法」も適用することも可能となる。

【0051】

また、この形態では、バイナリマスク作成部23で複数通りのバイナリマスクを作成し、限定混合信号抽出部24で、複数のバイナリマスクを用い、複数種類の限定混合信号を抽出することとした。これにより、1つの限定混合信号を分離するだけでは得られなかった分離信号が、他の限定混合信号を分離することによって得られるようになる。その結果、必要な原信号に対応する全ての分離信号を得ることが可能となる。

さらに、この形態の例では、第1の平均値及び第2の平均値となる平均値の組み合わせを変更することにより、複数通りのバイナリマスクを作成することとした。これにより、異なる原信号の組み合わせからなる複数の限定混合信号を得ることが可能となる。その結果、上記のように、必要な原信号に対応する全ての分離信号を得ることが可能となる。

また、この形態の例では、第1の平均値と第2の平均値との間に含まれない平均値を、新たな第1の平均値或いは第2の平均値として選択していくことにより、複数通りのバイナリマスクを作成することとした。これにより、異なる原信号の組み合わせからなる複数の限定混合信号を確実に得ることができる。その結果、上記のように、必要な原信号に対応する全ての分離信号を得ることが可能となる。

【0052】

なお、これまで説明してきたように、第1の実施の形態から第3の実施の形態におけるブラインド信号分離装置の構成はコンピュータによって実現することができる。この場合、各装置が有すべき機能の処理内容はプログラムによって記述され、このプログラムをコンピュータで実行することにより、上記処理機能をコンピュータ上で実現することができる。

また、この処理内容を記述したプログラムは、コンピュータで読み取り可能な記録媒体に記録しておくことができる。コンピュータで読み取り可能な記録媒体としては、例えば、磁気記録装置、光ディスク、光磁気記録媒体、半導体メモリ等のようなものでもよいが、具体的には、例えば、磁気記録装置として、ハードディスク装置、フレキシブルディスク、磁気テープ等を、光ディスクとして、DVD (Digital Versatile Disc)、DVD-RAM (Random Access Memory)、CD-ROM (Compact Disc Read Only Memory)、CD-R (Recordable) / RW (ReWritable) 等を、光磁気記録媒体として、MO (Magnet-Optical disc) 等を、半導体メモリとしてEEPROM (Electrically Erasable and Programmable-Read Only Memory) 等を用いることができる。

【0053】

また、このプログラムの流通は、例えば、そのプログラムを記録したDVD、CD-ROM等の可搬型記録媒体を販売、譲渡、貸与等することによって行う。さらに、このプログラムをサーバコンピュータの記憶装置に格納しておき、ネットワークを介して、サーバコンピュータから他のコンピュータにそのプログラムを転送することにより、このプログラムを流通させる構成としてもよい。

このようなプログラムを実行するコンピュータは、例えば、まず、可搬型記録媒体に記録されたプログラムもしくはサーバコンピュータから転送されたプログラムを、一旦、自己の記憶装置に格納する。そして、処理の実行時、このコンピュータは、自己の記録媒体に格納されたプログラムを読み取り、読み取ったプログラムに従った処理を実行する。また、このプログラムの別の実行形態として、コンピュータが可搬型記録媒体から直接プログラムを読み取り、そのプログラムに従った処理を実行することとしてもよく、さらに、このコンピュータにサーバコンピュータからプログラムが転送されるたびに、逐次、受け取ったプログラムに従った処理を実行することとしてもよい。また、サーバコンピュータから、このコンピュータへのプログラムの転送は行わず、その実行指示と結果取得のみによって処理機能を実現する、いわゆるASP (Application Service Provider) 型のサービスによって、上述の処理を実行する構成としてもよい。なお、上述した実施形態における

プログラムには、電子計算機による処理の用に供する情報であってプログラムに準ずるもの（コンピュータに対する直接の指令ではないがコンピュータの処理を規定する性質を有するデータ等）をも含むものとする。

【0054】

また、この発明は上述の実施の形態に限定されるものではない。例えば、上述の各形態ではコンピュータ上で所定のプログラムを実行させることにより、本装置を構成することとしたが、これらの処理内容の少なくとも一部をハードウェア的に実現することとしてもよい。また、上述の各種の処理は、記載に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的にあるいは個別に実行されてもよい。その他、本発明の趣旨を逸脱しない範囲で適宜変更が可能であることはいうまでもない。

【産業上の利用可能性】

【0055】

本技術により、例えば、様々なノイズ・妨害信号が存在する環境下においても、目的信号を精度よく分離抽出することが可能となる。例えば、オーディオ分野に応用した場合、音声認識機の入力マイクロホンと話者が離れた位置にあり、マイクロホンが目的話者音声以外の音まで集音してしまうような状況でも、目的音声を分離抽出することにより、認識率の高い音声認識系を構築することができる。

【図面の簡単な説明】

【0056】

【図1】 ブラインド信号分離装置の全体構成を例示したブロック図。

【図2】 信号選択部及び信号分離部の詳細構成を例示したブロック図。

【図3】 ブラインド信号分離方法を説明するためのフローチャートの例示。

【図4】 相対値の分布図及びバイナリマスクを例示した図。

【図5】 信号選択部と信号分離部の構成を例示したブロック図。

【図6】 ブラインド信号分離方法を説明するためのフローチャートの例示。

【図7】 バイナリマスクを例示した図。

【図8】 (a) は、このブラインド信号分離技術を概念的に例示したブロック図であり、(b) は、 $N=M=2$ の場合における独立成分分析による分離過程を例示したブロック図である。

【図9】 スパース性によるブラインド信号分離のブロック図。

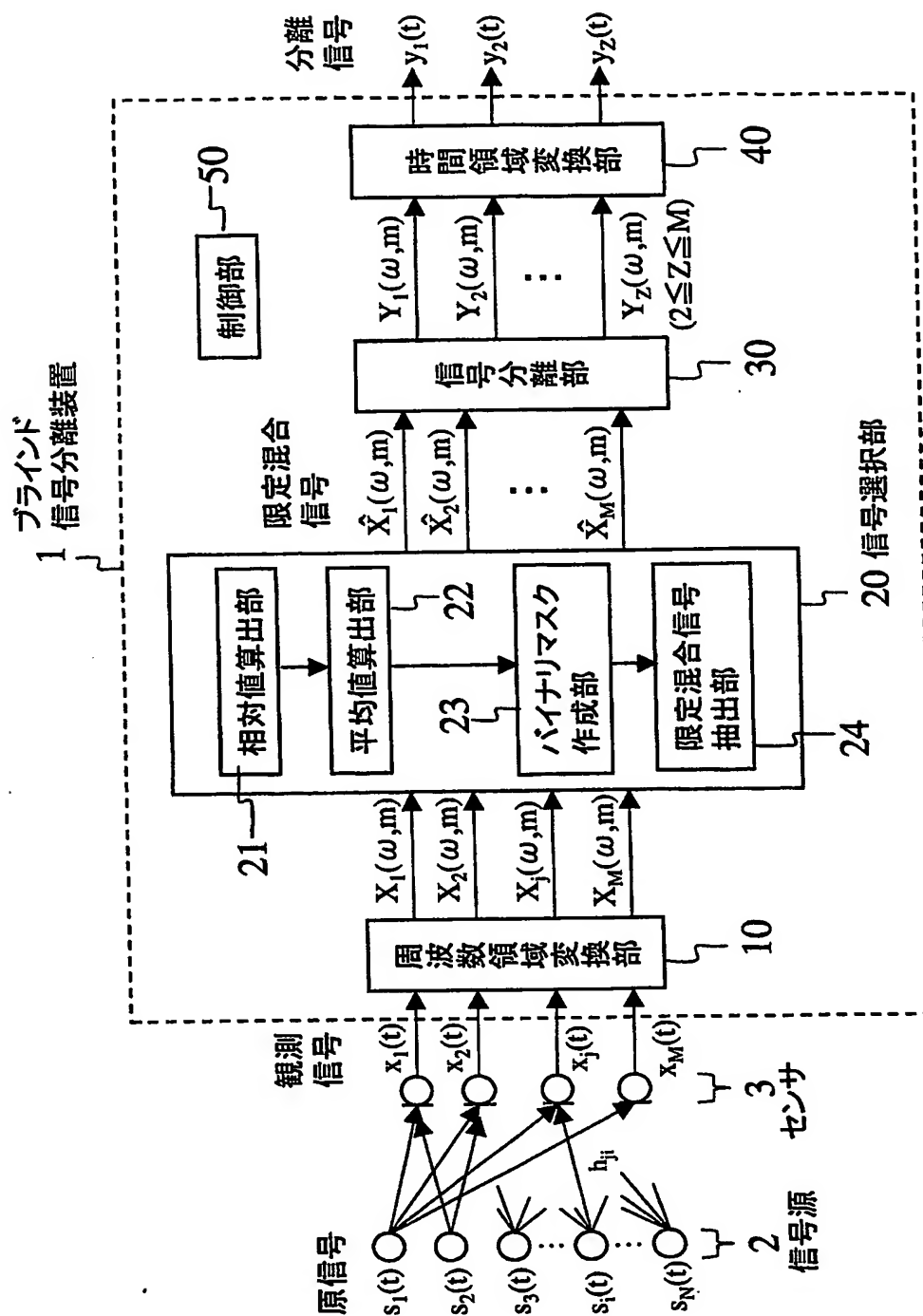
【図10】 位相差・振幅比の分布を例示した図。

【符号の説明】

【0057】

- 1 ブラインド信号分離装置
- 10 周波数領域変換部
- 21 相対値算出部
- 22 平均値算出部
- 23 バイナリマスク作成部
- 24 限定混合信号算出部
- 30 信号分離部
- 40 時間領域変換部

【書類名】 図面
【図 1】



【図 2】

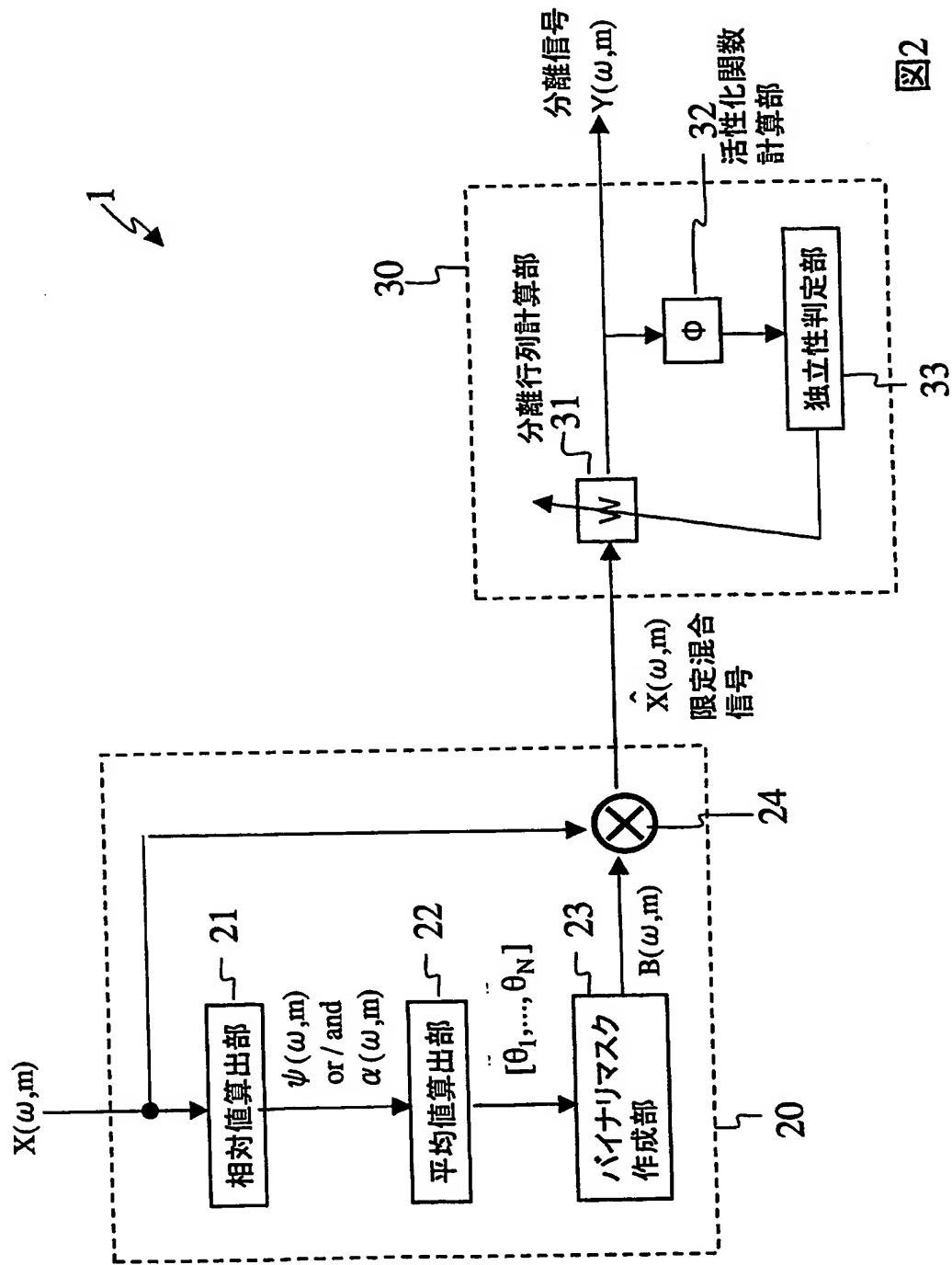


図2

【図3】

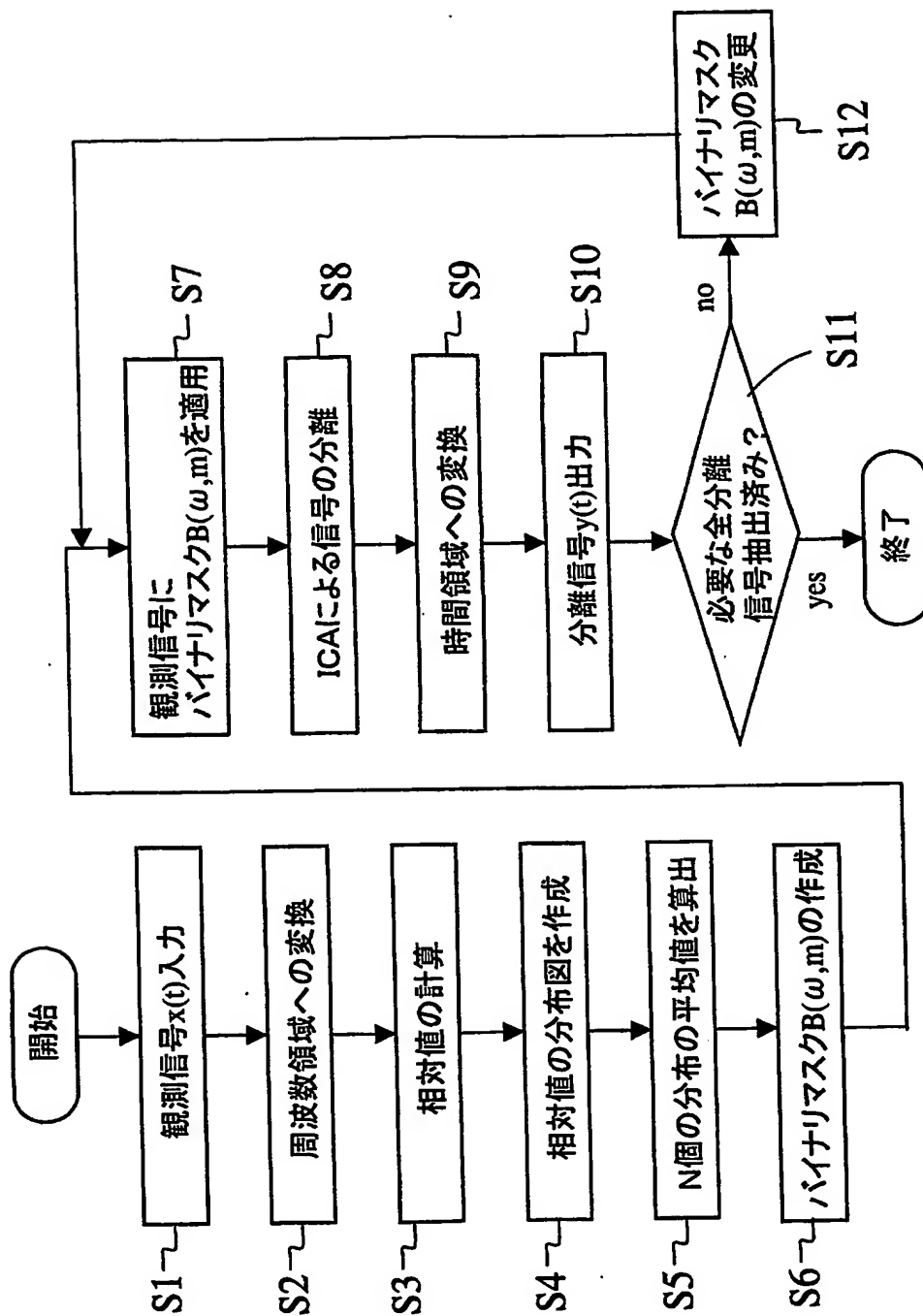


図3

【図4】

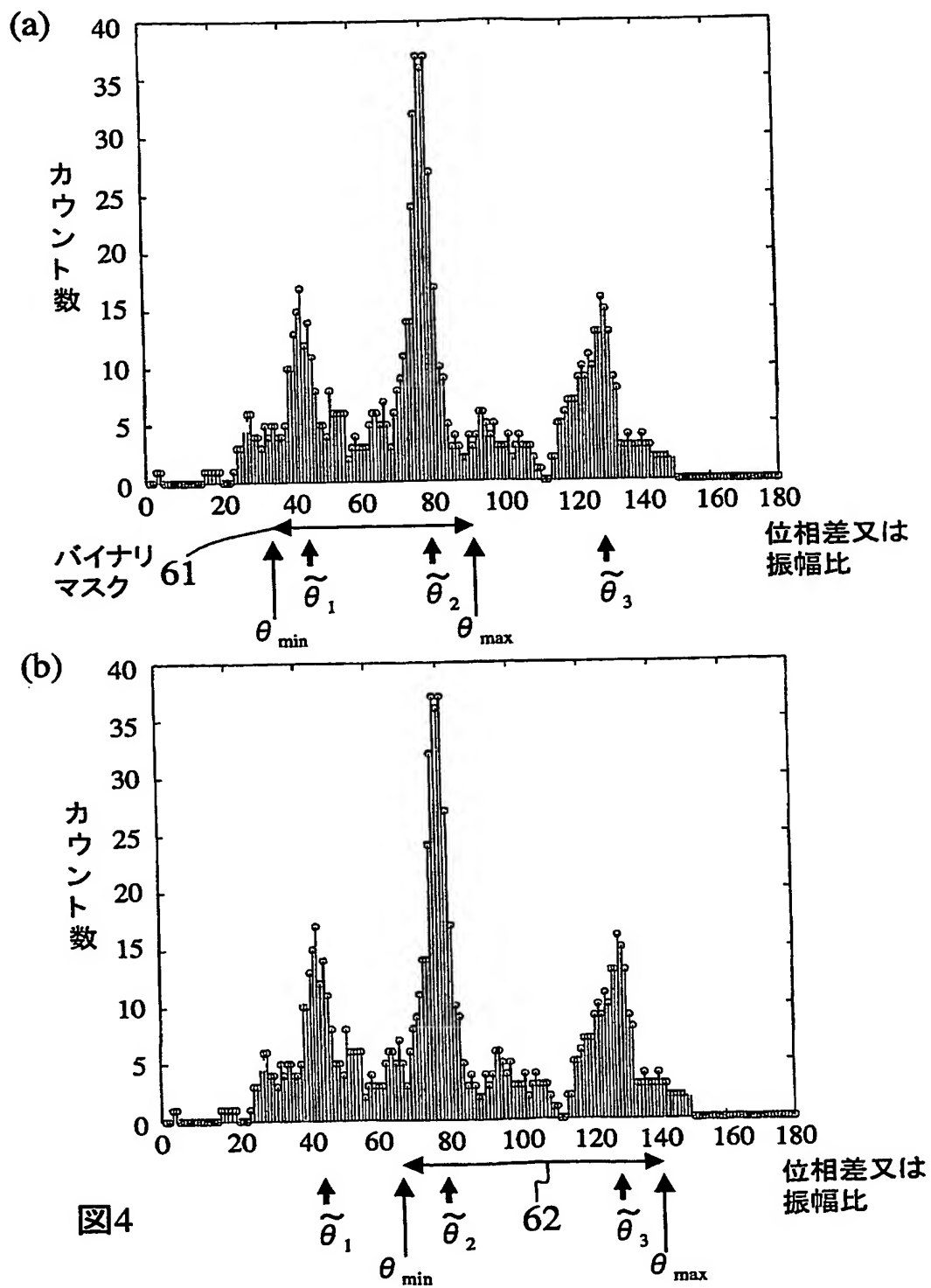


図4

【図 5】

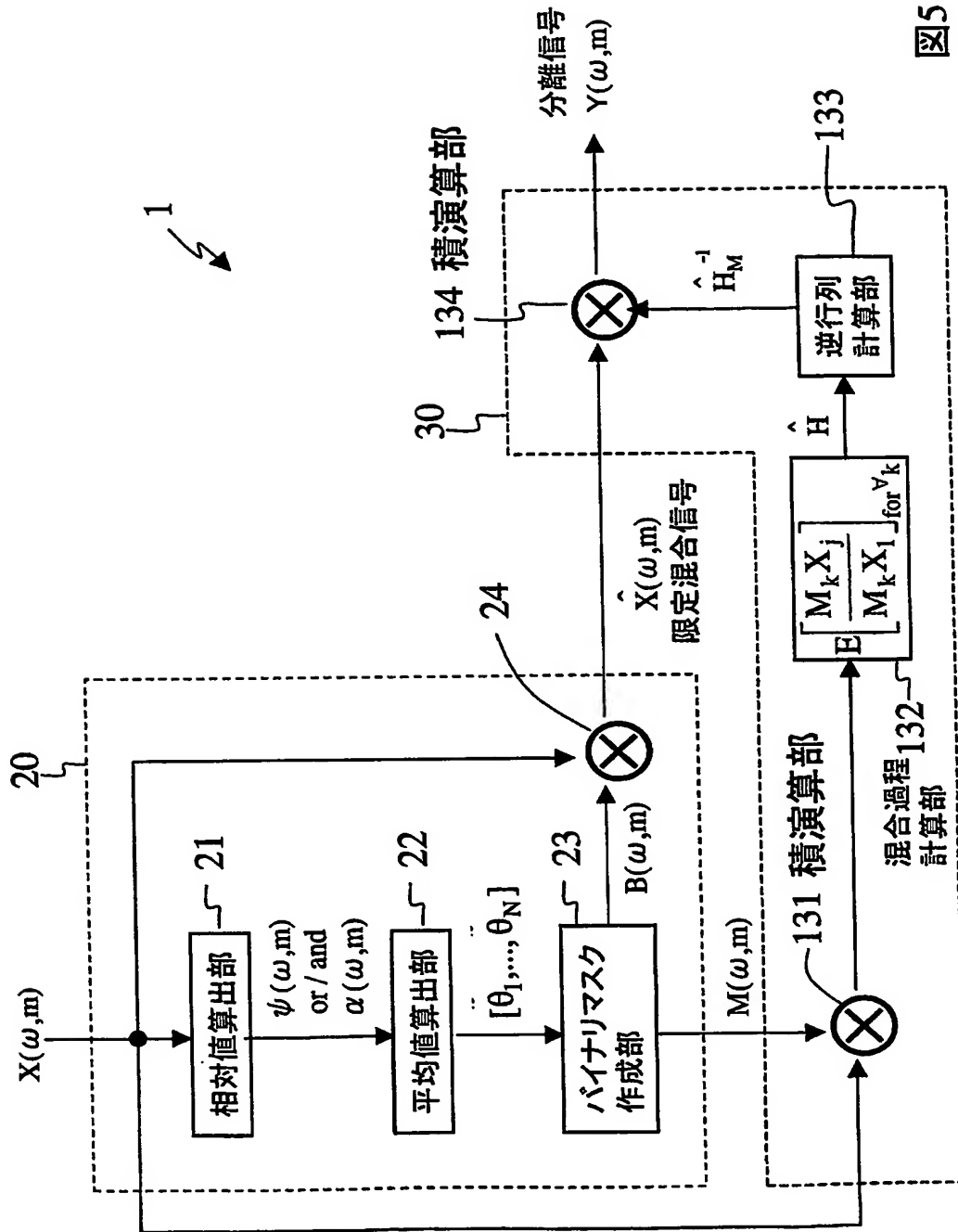


図5

【図 6】

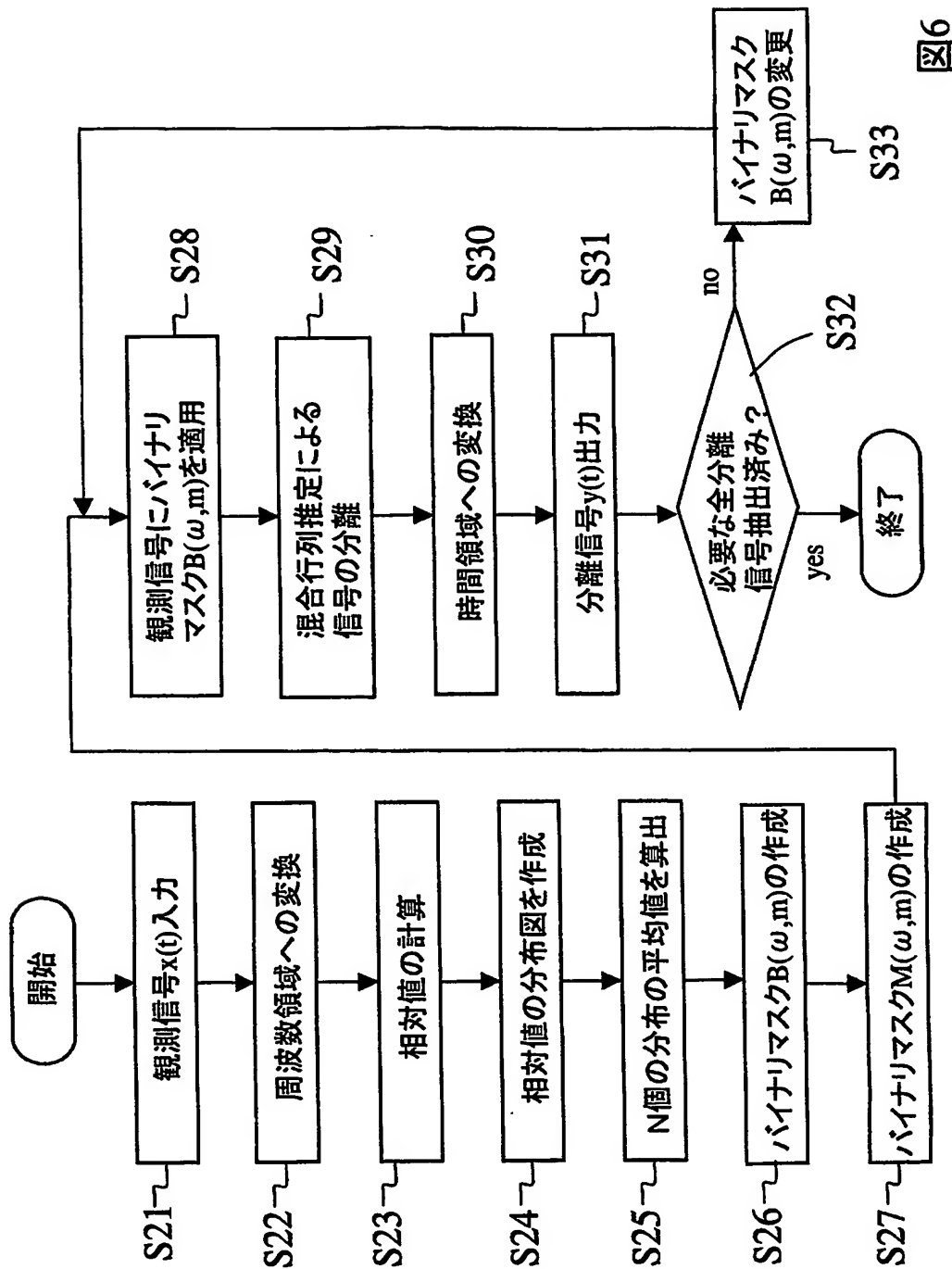


図6

【図 7】

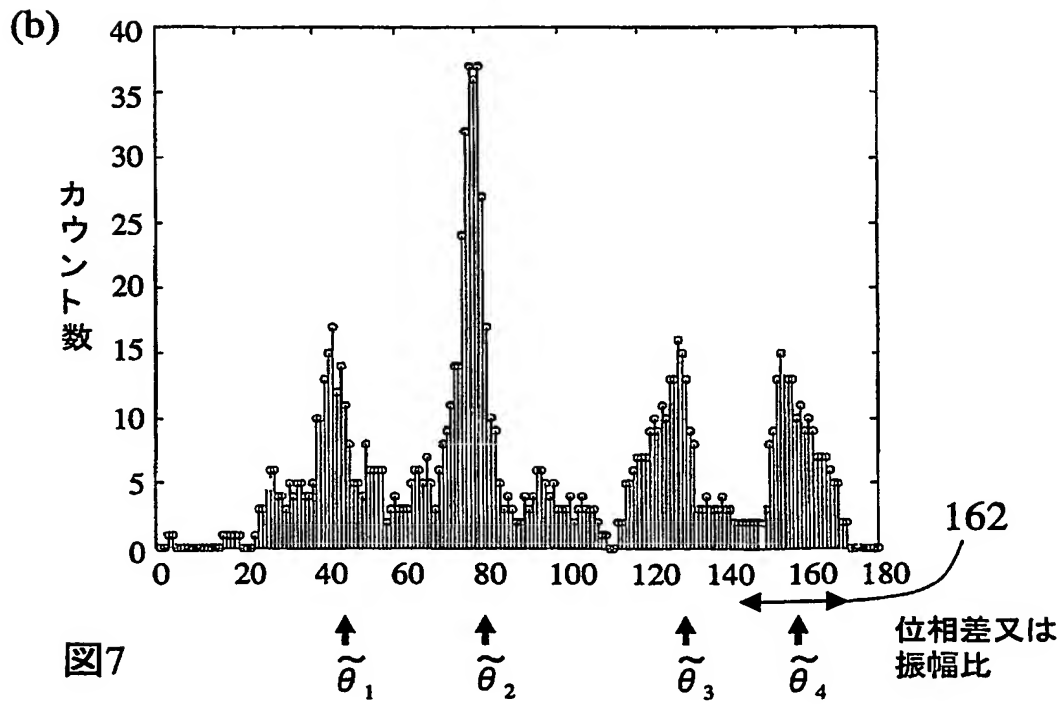
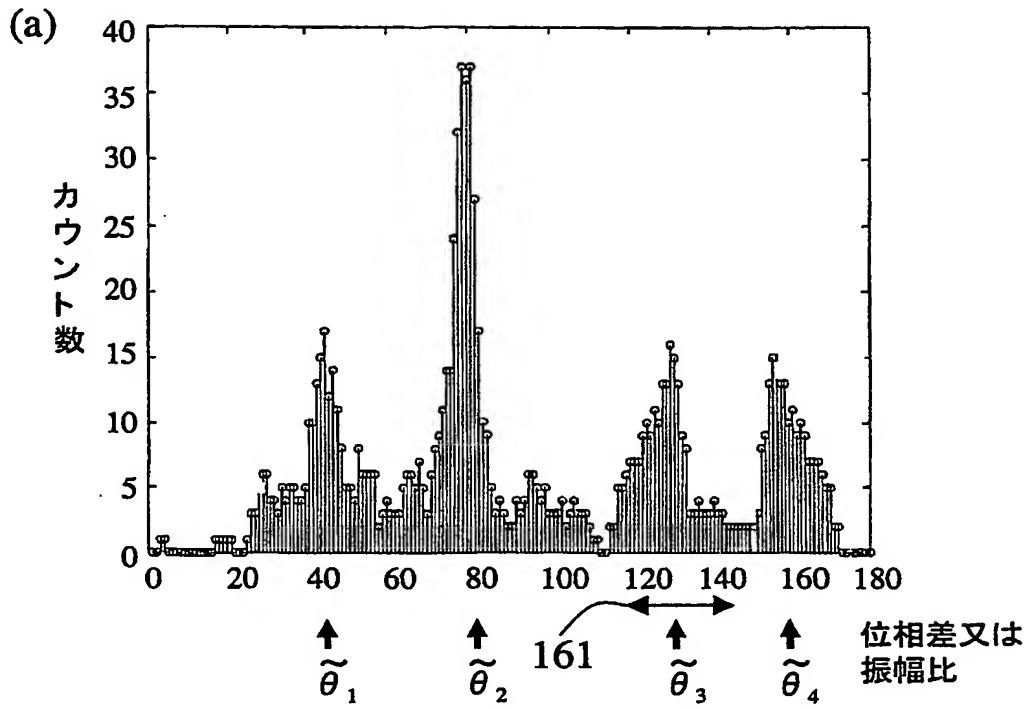
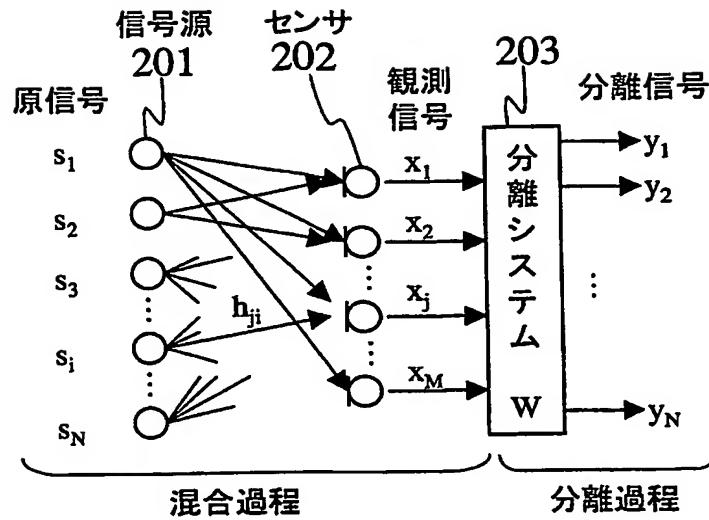


図7

【図8】

(a)



(b)

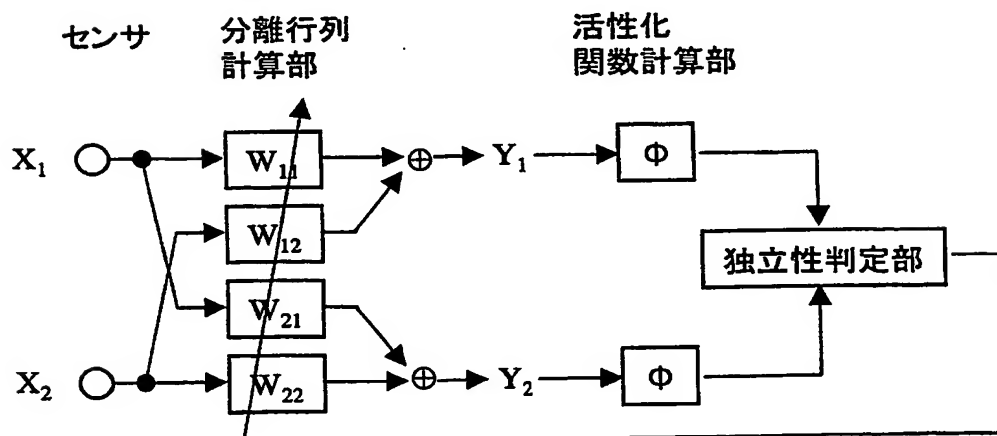


図8

【図9】

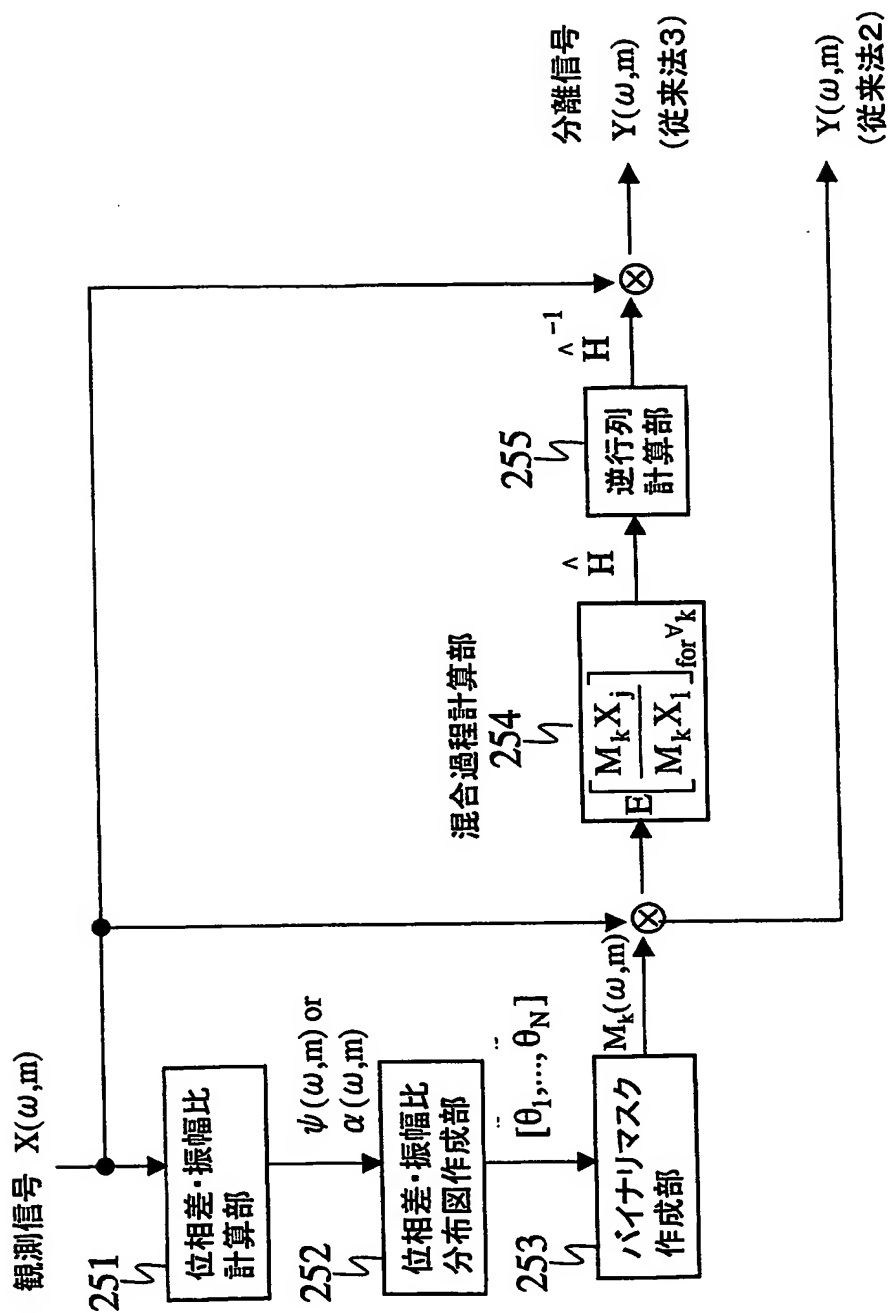


図9

【図10】

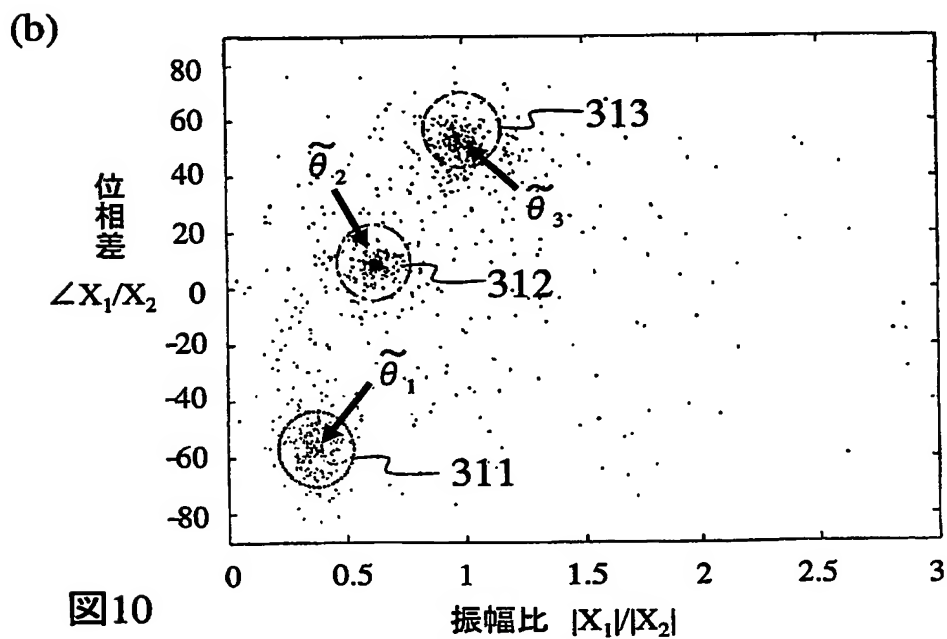
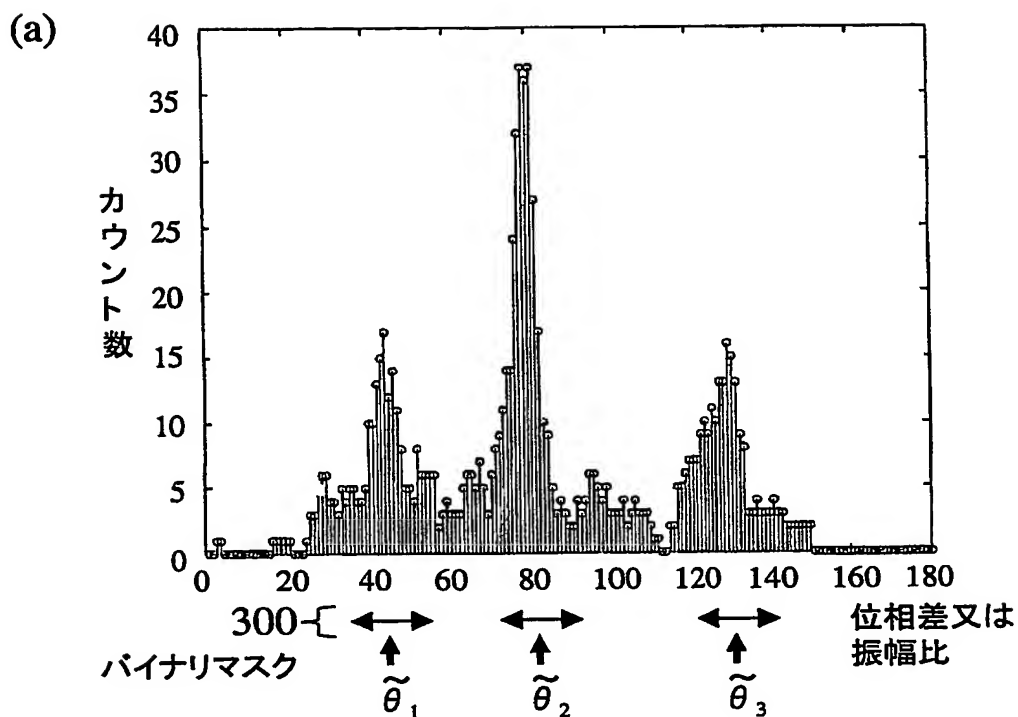


図10

【書類名】 要約書

【要約】

【課題】 信号源の数 N とセンサの数 M が $N > M$ の関係にある場合に、混合信号を高い品質で分離する。

【解決手段】 センサ 3 で観測した観測信号を、周波数領域変換部 1 0 で周波数領域の信号に変換し、相対値算出部 2 1 で、この周波数領域の信号を用い、センサ間における位相差等の相対値を各周波数について算出する。次に、算出した相対値の分布がピークとなる位置の相対値（平均値）を算出し、この平均値を用い、変換した周波数領域の信号から、2 個以上 M 個以下の信号源から発せられた信号によって構成される混合信号（限定混合信号）を抽出するためのバイナリマスクを作成する。そして、この作成したバイナリマスクを用い、限定混合信号を抽出する。そして、この限定混合信号に対し I C A 等の原理的に Musical Noise が起こらない信号分離手法を適用する。

【選択図】 図 1

特願 2 0 0 3 - 3 0 9 7 2 0

出 願 人 履 歴 情 報

識別番号

[0 0 0 0 0 4 2 2 6]

1. 変更年月日

1 9 9 9 年 7 月 1 5 日

[変更理由]

住所変更

住 所

東京都千代田区大手町二丁目 3 番 1 号

氏 名

日本電信電話株式会社